BIOS6640: Final R Project

Ana Babinec

# The WHO African Region and Malaria

The WHO African Region continues to bear the brunt of the global burden of malaria, despite the decline in malaria cases and malaria death rate that occurred during 2000 and 20151. In 2016, it is estimated that 445,000 people died from malaria worldwide, with the majority being children in sub-Saharan Africa5. In addition to the high death incidence, malaria causes socioeconomic struggles in population as it results in absenteeism from work and prevents students from attending school due to illness and symptoms. The biology and infection pathway of malaria is well understood—malaria is caused by the *Plasmodium* parasite—female *Anopheles mosquitoes* act as malaria vectors and the disease is spread to humans when one is bitten by the infected mosquito4. However, there is not as clear of an understanding as to why Africa continues to bear the brunt of the burden of malaria, despite increased funding, innovations and resource allocation.

Insufficient and inaccurate methods of malaria recording and reporting are part of the explanation—it wasn’t until 2000 that the United Nations dedicated a Millennium Development Goal to reducing malaria incidence5. This commitment led to improvements in health system infrastructure and malaria control – from 2000 to 2015 over 6.2 million deaths were averted and the global mortality rate and incidence rate declined by 58% and 37% respectively4. However, despite the regional efforts across Africa, discrepancies still remain in malaria burden across the continent.

# Prior research on spatial analysis of malaria

Prior research has utilized spatial analysis to understand the association between malaria incidence and spatio-temporal climate variations. Significant associations were found among incidence anomalies and climate patterns. Remote influences such as monsoon rainfalls and large-scale climate events such as the Indian Ocean Diploe and El Nino Southern Oscillation both have been linked to malaria incidence2. There is consensus among spatial analysis studies that temperature across Africa, as a whole, is a strong predictor for the transmission of *Plasmodium,* however there is disagreement as to whether rainfall or temperature is the primary driver of malaria incidence4. Studies have incorporated lag time between climate variables and malaria incidence to better explain the relationship as malaria incidence is dependent on the life cycle of the mosquito vector and parasite and the number of days between infection and diagnosis. Specifically, a study focused on the seasonally lagged effects of climatic factors on malaria incidence in South Africa found that malaria incidence rate anomalies were associated with higher than normal precipitation in southern Mozambique at a lag of two months1.

# Motivation and objective

Despite the numerous studies on spatial analysis of climatic variables and malaria incidence, there is inconclusive evidence of the statistical significance of lagged temperature and precipitation on the malaria incidence. Additionally, previous studies have been limited by the lack of cases and data surrounding climatic variables as well the aggregation of malaria incidence to high and low rates, rather than raw values. In regards to our country of interest, Mozambique, previous spatial analysis studies on malaria had to impute climatic data due to incompleteness. In addition, prevalence data was used spanning from 1999-2008—the use of data prior to the establishment of the MDG of 2000 may affect the analysis as commitment to decreasing incidence came after the UN declared malaria an MDG.

Given that Mozambique still has high year-round malaria incidence it is important that a thorough spatial analysis of the association between climatic variables and malaria incidence is performed to better understand the ology of malaria. The goal of the following analysis is to better describe how the temporal and spatial variation relate to the variation in malaria incidence. Specifically, analysis focused on 1) the impact of weekly rainfall and temperature on malaria incidence and 2) the incorporation of lagged variables for weather data and their relative significance.

# Data and Analysis Methods

Malaria data

Malaria case data for Mozambique was collected from 2010-2016. Data was recorded every epi-week—an epiweek begins on a Sunday and end on a Saturday, with the first epi-week of the year ending on the first Saturday of January. District level data, including total population, proportion of the total population under 5 years of age, the number of malaria cases, and a unique identifier for district are all present in the data set. Malaria incidence in cases per 1,000 population in children under 5 years of age was calculated using the under 5 weight and total province population.

Explanatory data

Weather data was also included in the dataset, and was also collected by epi-week from 2010-2016 by province. Data includes weekly average rainfall, weekly total rainfall, weekly average temperature, relative humidity, days above a threshold temperature and threshold precipitation, as well as the surface barometric pressure (which is a general indicator of large-scale weather activity and exhibits a strong seasonal cycle). Weekly averages by province will be used in the analysis to match the week units of malaria cases.

# Analysis Plan

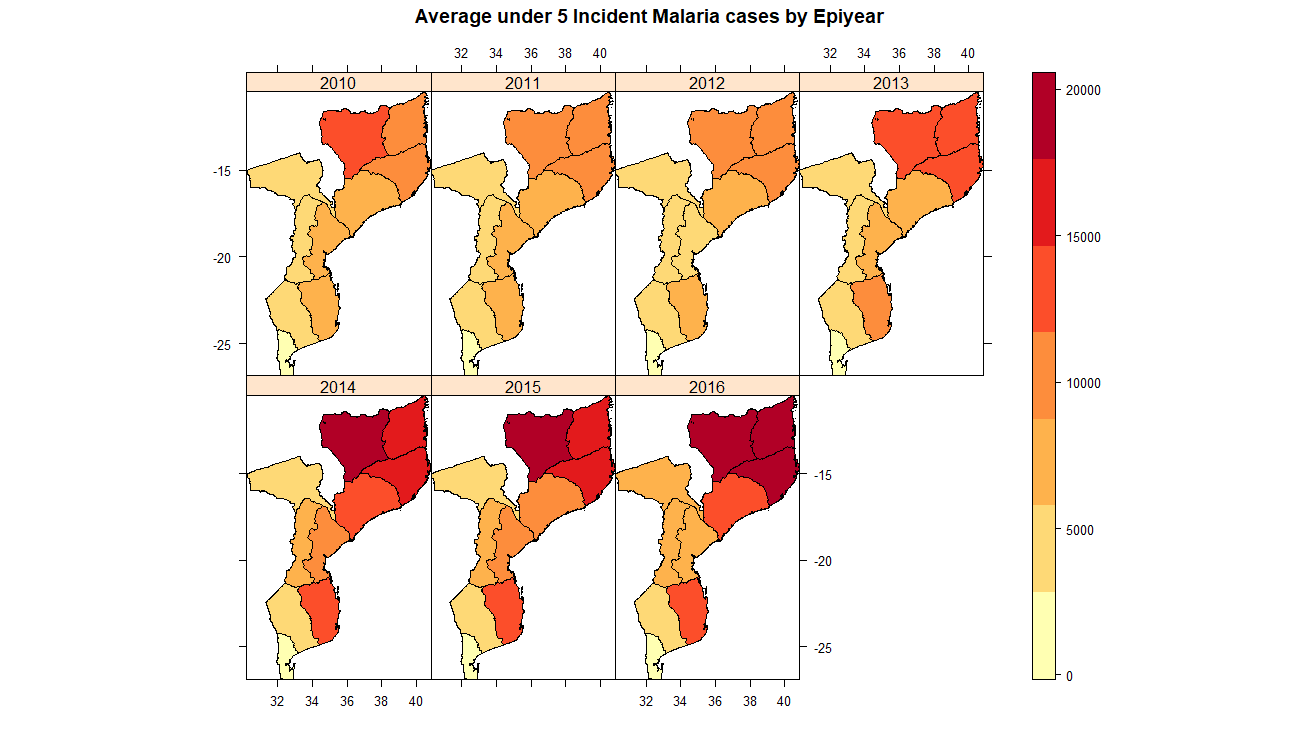
For describing the temporal and spatial variation in the data and to draw conclusions about the relationships between the variables, a spatial analysis was performed, incorporating lagged effect of weather at 0, 2,4, and 8 weeks, on malaria incidence per 1,000 population in children under 5.

Elementary analysis was performed to understand the current distribution of malaria across Mozambique as well as the current weather and temporal patterns across the country. Province level differences were determined and recorded, for both malaria incidence and weather variables. In addition to across province differences, changes in incidence and weather conditions were analyzed across years.

Analysis was performed on a data set obtained from Dr. Colborn, University of Colorado-Anschutz Medical Campus using RStudio Version 3.5.1.

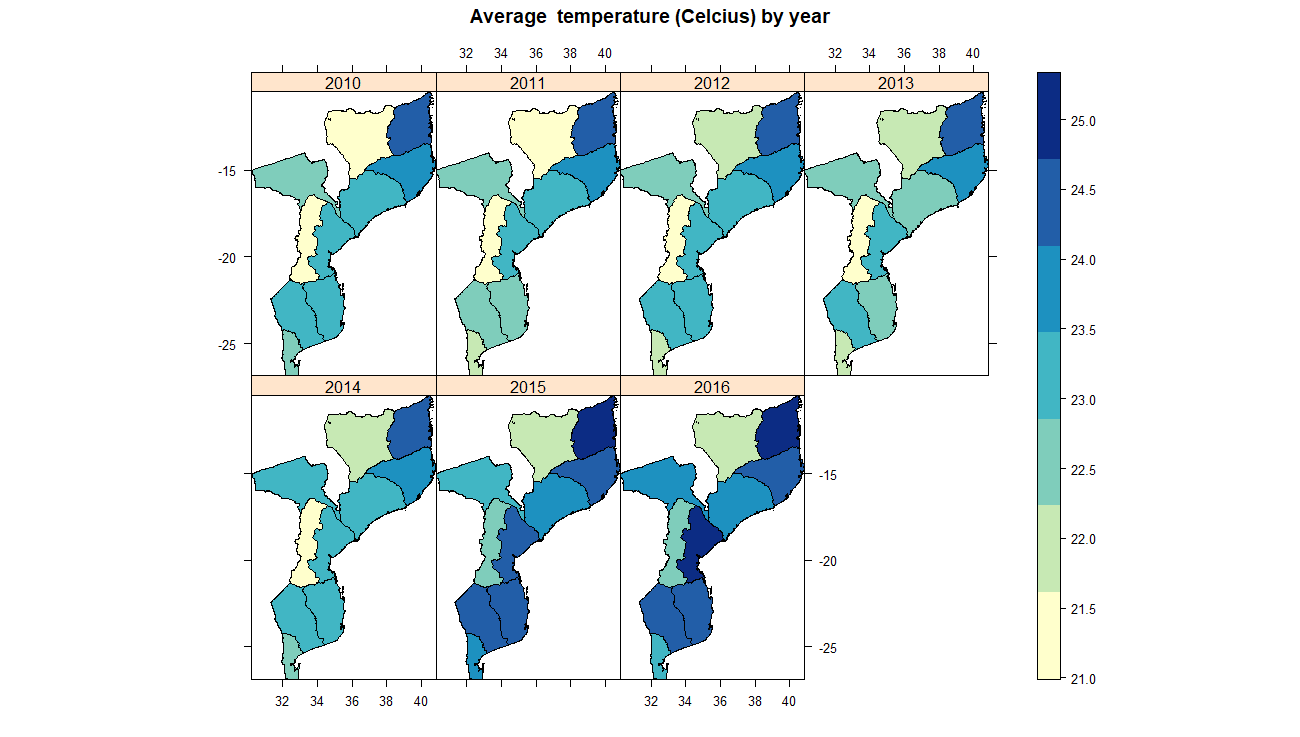
# Results

Average under 5 malaria incidence across Mozambique

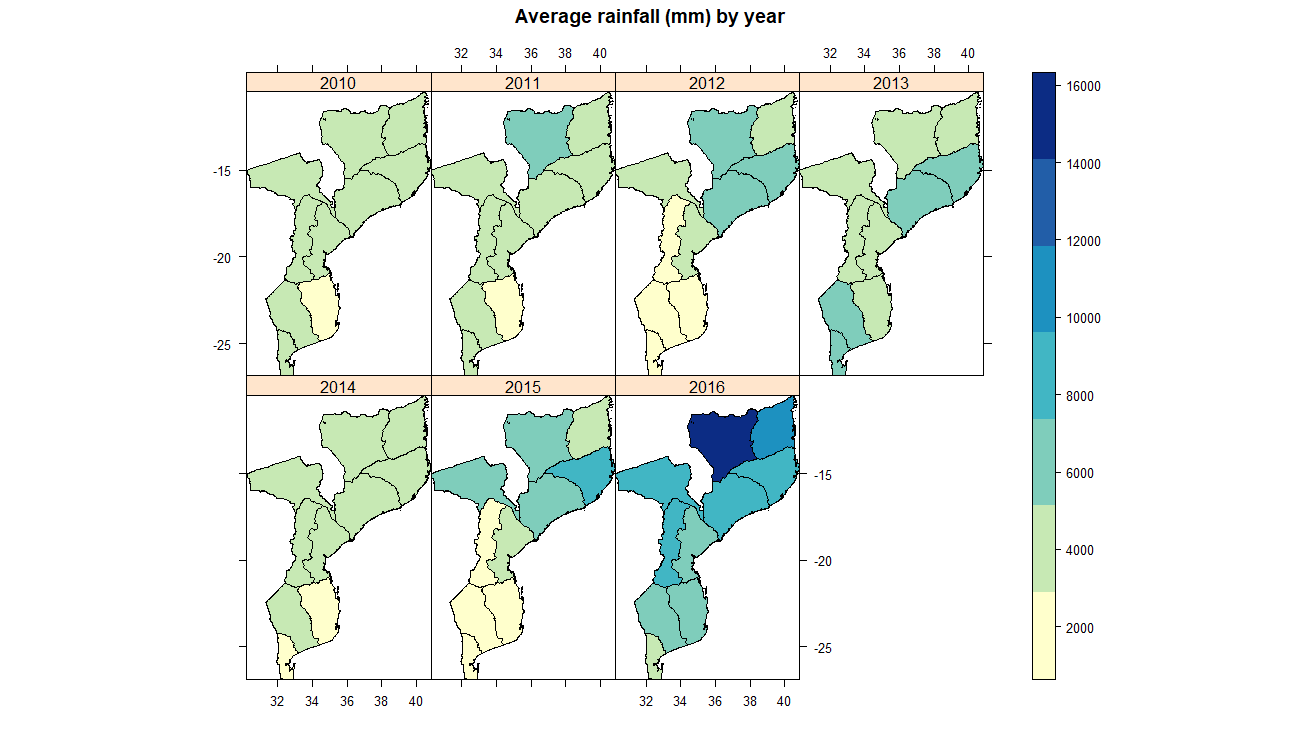
Based on the figure to the right (Figure 1), it is clear that malaria incidence in children under 5 is increasing over time. The Northern region experienced the greatest increase in malaria incidence, whereas the Center region experienced the least increase in malaria incidence. Change in malaria incidence does seem to depend on geographic region, which could be due to differences in weather conditions, such as temperature and rainfall—both known influences of malaria. In 2010, six provinces had a malaria incidence under 10 cases/1,000, whereas only three provinces had malaria incidence under 10 cases/1,000 in 2016. However, while cases are more prevalent in the Northern and Center region, cases aren’t noticeable clustered in these regions—across Mozambique theres is significant incident malaria cases.

**Figure 1** Under 5 malaria incidence cases by province across Epiyear.

## Explanatory variables (rainfall and average temperature) across Mozambique



**Figure 2** Average temperature by province across Epiyear.



**Figure 3** Average rainfall (mm) by province across Epiyear.

The plots above show the distribution of average weekly temperature (oC) and average weekly rainfall (mm) across Mozambique. From the plots it is clear that rainfall is variable by year across all provinces, with the northern region experiencing the greatest rainfall across nearly all years. Temperature is more variable across Mozambique, however there has been a general increases over time. The northern and coastal region have increased in temperature the most. The correlates with the increase in malaria incidence seen in Figure 1.

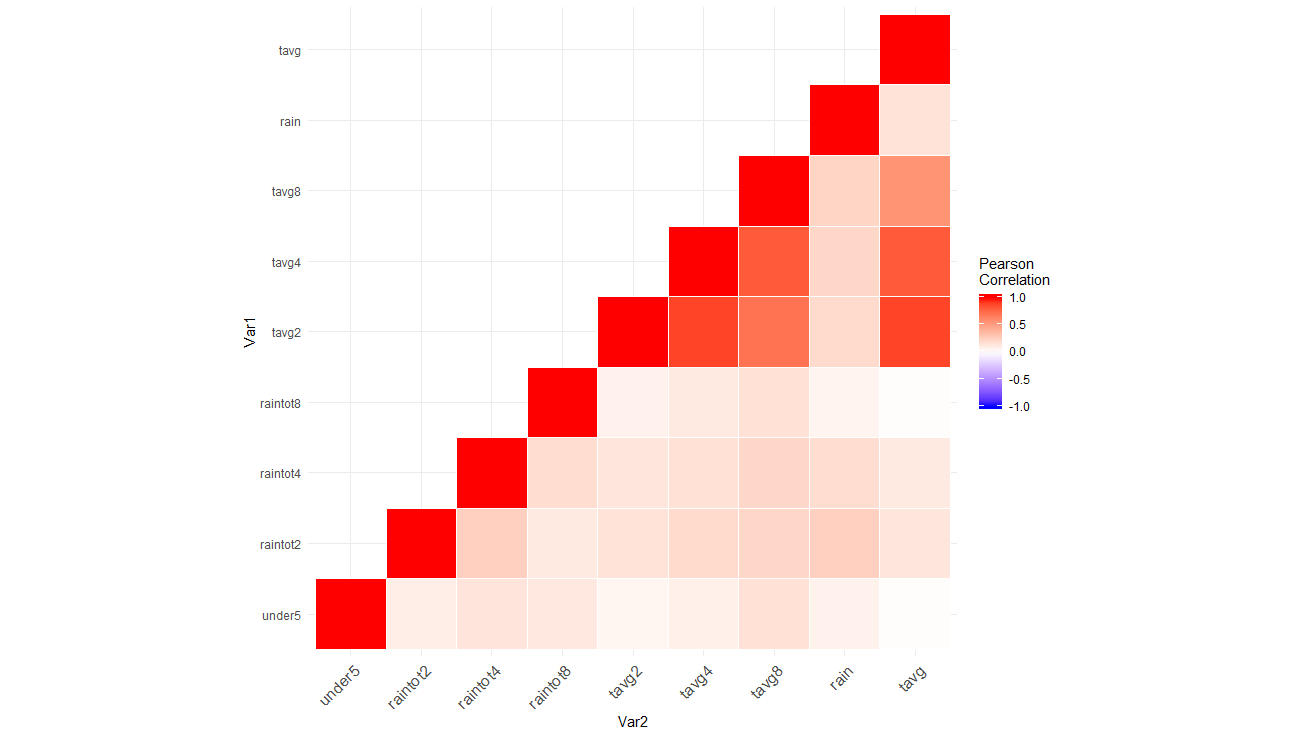
## Effect of weather on malaria incidence

Average rainfall (mm) and temperature (oC)

Based on the analysis of un-lagged weekly rainfall by Epiyear and region with under 5 malaria incidence, it is apparent that lagged effect does exist, as rainfall is lowest during Epiweeks 20-40 across all regions, however malaria incidence drops to its near minimum value at week 30, which is consistent across all regions. The rainfall patterns in the Coastal and Southern regions are less variable than the Northern and Center region, which is consistent with the previously discussed higher malaria incidence in these regions.

With average temperature, a similar pattern exists, as average temperature increased (up to around 26oC) incidence of malaria trended upward. After this temperature, incidence decline. Rainfall follows a similar trajectory of malaria incidence, there is an increase in rainfall up until 25oC, and then a sharp decline. To further explain the relationship between temperature and rainfall with malaria incidence, lagged-effects were created.

# Correlation of lagged rainfall and average temperature on incidence

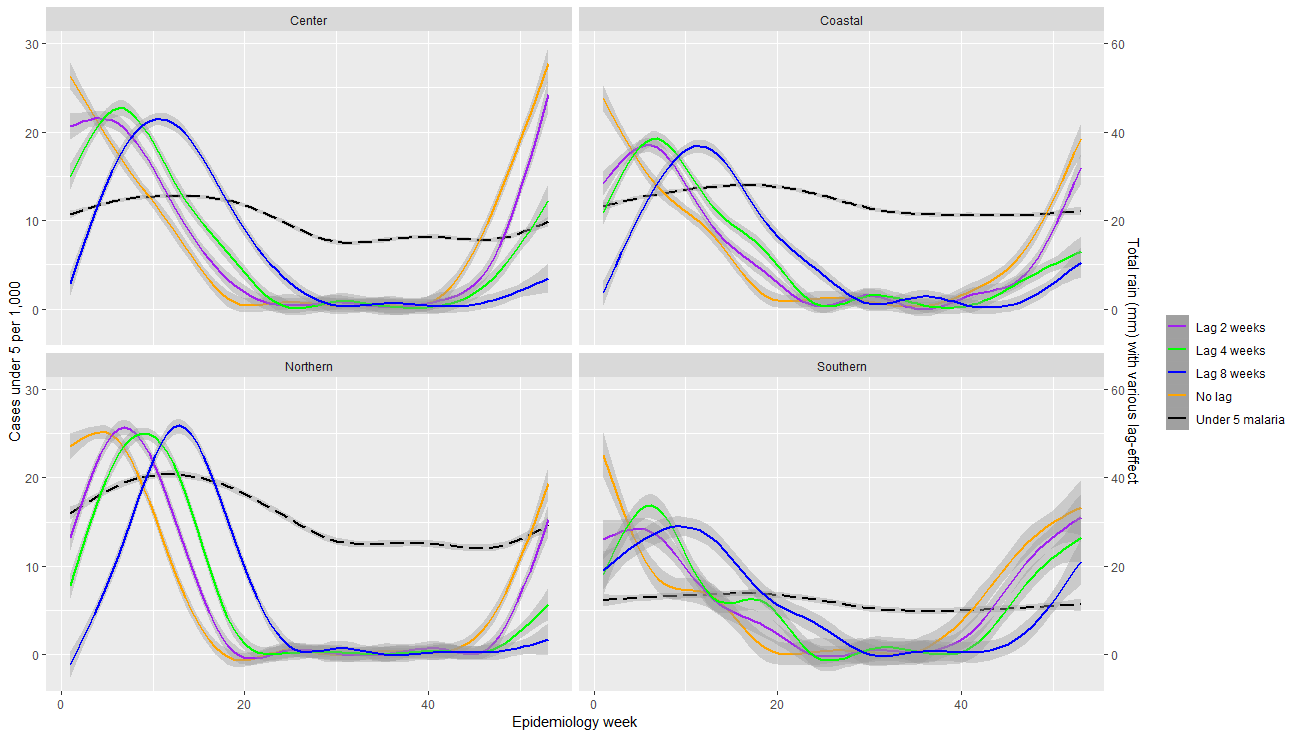


**Figure 4** Correlation plot of lagged weather variables with under 5 malaria incidence.

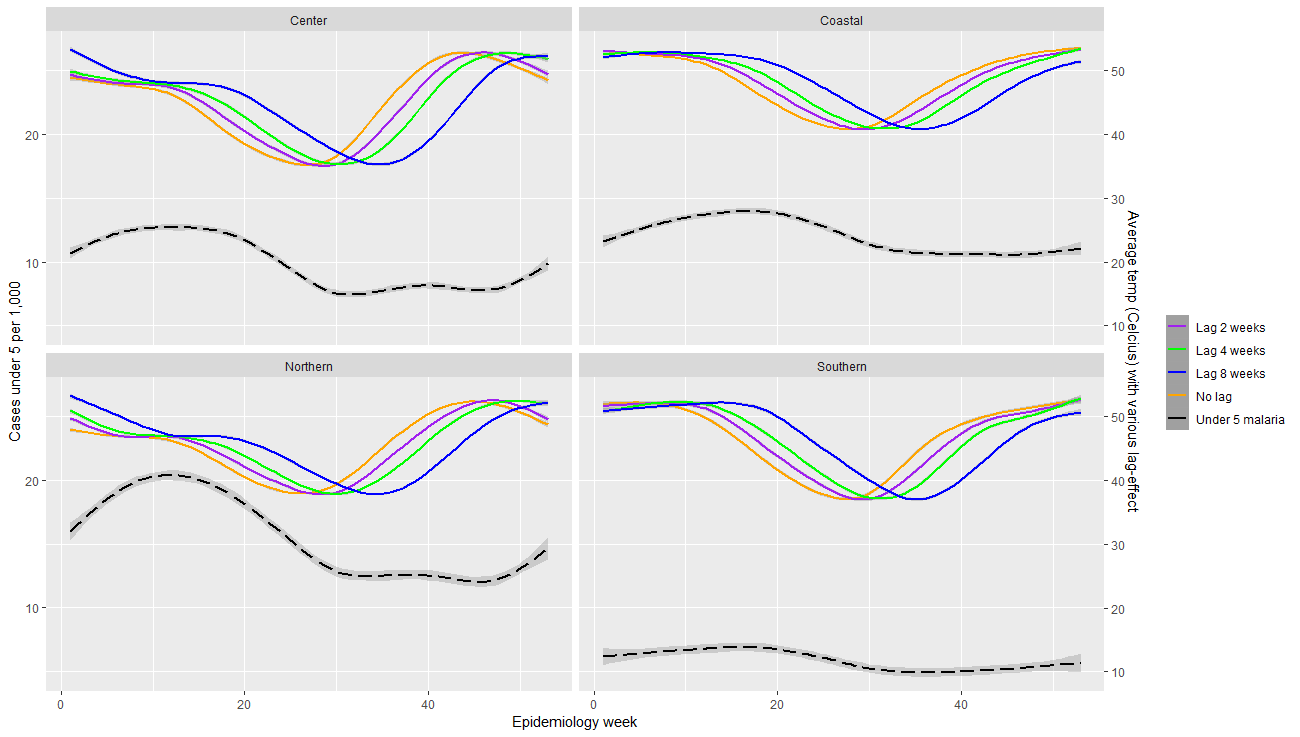
From the correlation plot (right) it is apparent that the lagged effects of 4 weeks for rainfall and 8 weeks for temperature are most correlated with malaria incidence in children under 5. This is consistent with the smoothed plots generated (below) which plots the incidence cases under 5 in each region by epidemiology week with the total rainfall (mm) with lagged effect of 0, 2,4, and 8 week and the average temperature (Celsius). The four week lag varies consistently with malaria incidence across regions and years. There is a considerable improvement over ‘No lag’, further supporting the delayed effect of climatic variables and the infection and incubation period of malaria. With regards to lagged temperature effect, an 8 week lag is similarly more consistent with the change in malaria incidence over Epiweeks. Lags of 2 and 4 weeks resulted in very similar trends as ‘no lag’, which, again, further supports the delayed effect of climatic variables.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | Under5 | rainTot | rainTot2 | rainTot4 | rainTot8 | Tavg | Tavg2 | Tavg4 | Tavg8 |
| Under5 | 1.0 | 0.07 | 0.09 | **0.14** | 0.12 | 0.01 | 0.05 | 0.08 | 0.16 |

# **Table 1:** *Correlation of lagged variables (rain and avg. temp) with under 5 malaria incidence. Week-lag denoted in variable name.*



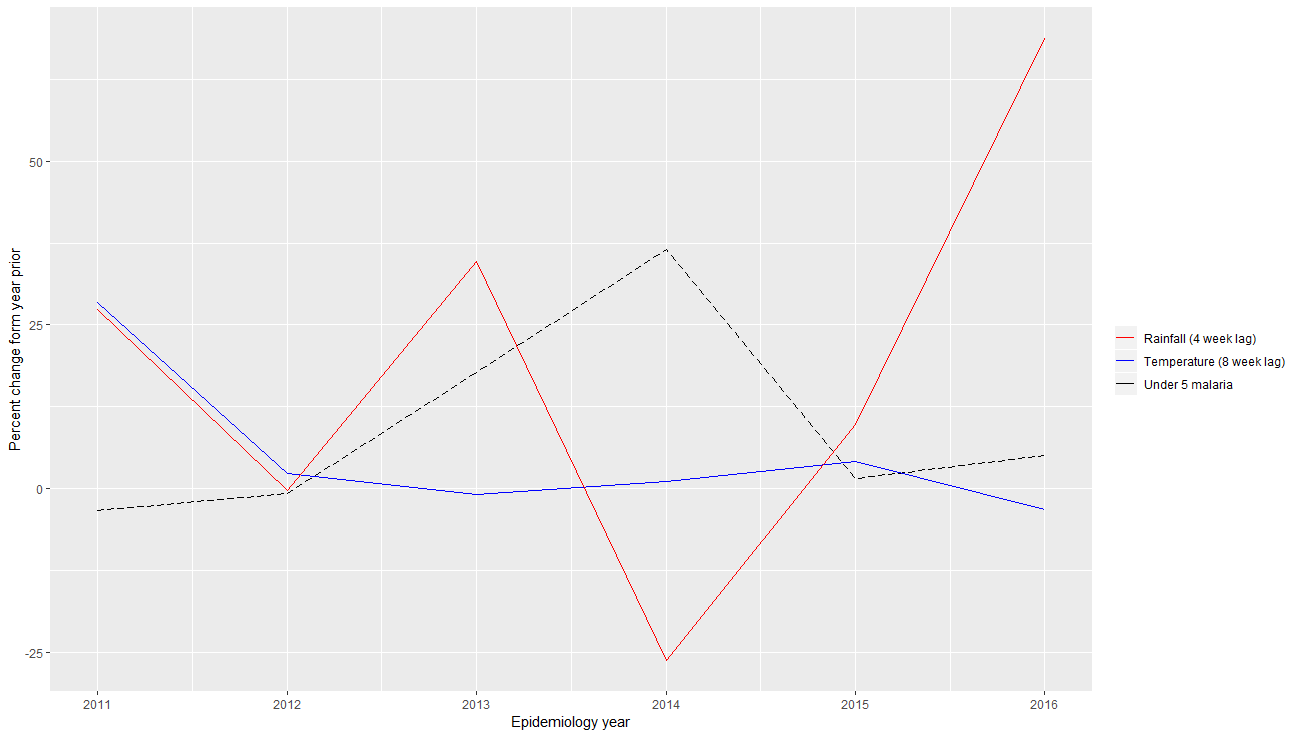
**Figure 5** Total rain (mm) with various lag effect versus under 5 incidence by Epiweek and Region.



**Figure 6** Average temp (Celsius) with various lag effect versus under 5 incidence by Epiweek and Region.

# Change in lagged effects over years

Taking into account the correlation between 8 week lagged temperature, 4 week lagged rain and malaria incidence in children under 5, analysis was performed to determine the percent change in these variables between Epiyears. The plot below shows that there is the most variability in rainfall and malaria incidence—rainfall and temperature had similar percent changes up until 2012 (i.e. percent change from 2011 to 2012). Temperature changes rather consistently across years, indicating it is less effected by climatic events such as La Nina. Thus, there is reason to believe there is higher correlation between rainfall and malaria incidence, and efforts should be directed towards controlling the increase in incidence increased rainfall brings.



**Figure 7** Percent change in lagged weather effects by Epiyear.

# Discussion and further analysis

With this analysis and prior research, it is clear that there are lagged effects of weather on malaria incidence as well as between region variability in Mozambique. Results from the smoothed plots of the lagged effects of climate on under 5 incidence as well as the correlation plot, show that there are significant associations between malaria and weather variables. While these associations are present across all regions of Mozambique, some regions have higher underlying malaria incident rates, as well as rainfall and average temperature, which are seen from the maps of rainfall, temperature and malaria. However, there are some limitations in this analysis—no statistical model was developed to analyze statistical correlation between the measures of interest—this model would incorporate random effects (for province) and allow for the interaction between weather variables. In addition, societal and political factors such as current political state, healthcare delivery infrastructure, global commitment to malaria and perception of malaria risk among community members are potential confounders, and were not addressed in this analysis. Further analysis should investigate such a model and allow for nonlinear trends in weather effects. Additionally, geographically weighted regression and ‘nearest neighbor interpolation’ should be explored to determine how one province influence those provinces it neighbors.

# References

1. Ferrão, João Luís, et al. “Malaria Mortality Characterization and the Relationship between Malaria Mortality and Climate in Chimoio, Mozambique.” *Malaria Journal*, vol. 16, no. 1, 2017, doi:10.1186/s12936-017-1866-0.

2. Ikeda, Takayoshi, et al. “Seasonally Lagged Effects of Climatic Factors on Malaria Incidence in South Africa.” *Scientific Reports*, vol. 7, no. 1, 2017, doi:10.1038/s41598-017-02680-6.

3. K. Colborn, Exploratory Data Analysis course notes. *See /ExploratoryDataAnalysisLec%20(1).html (2018).*

4. Noor, Abdisalan M, et al. “The Changing Risk of Plasmodium Falciparum Malaria Infection in Africa: 2000–10: a Spatial and Temporal Analysis of Transmission Intensity.” *The Lancet*, vol. 383, no. 9930, 2014, pp. 1739–1747., doi:10.1016/s0140-6736(13)62566-0.

5. World Health Organization. “Malaria.” *World Health Organization*, World Health Organization, afro.who.int/health-topics/malaria.

Code available at: https://github.com/aebabinec/BIOS6640\_Bab

Collaborated with Yuli Chen and Anjin Singh.